

A Limit Formula for Joint Spectral Radius with p -radius of Probability Distributions

M. Ogura^{a,*}, C. F. Martin^a

^a*Department of Mathematics and Statistics, Texas Tech University, Broadway and Boston, Lubbock, TX 79409-1042, USA*

Abstract

In this paper we show a characterization of the joint spectral radius of a set of matrices as the limit of the p -radius of an associated probability distribution when p tends to ∞ . Allowing the set to have infinitely many matrices, the obtained formula extends the results in the literature. Based on the formula, we then present a novel characterization of the stability of switched linear systems for an arbitrary switching signal via the existence of stochastic Lyapunov functions of any higher degrees. Numerical examples are presented to illustrate the results.

Keywords: Joint spectral radius, p -radius, Lyapunov functions, absolute exponential stability

2000 MSC: 15A60, 15A48, 93D05, 93E15

1. Introduction

The joint spectral radius of a set of matrices, originally introduced in the short note [1], is a natural extension of the spectral radius of a single matrix and has found various applications in, for example, wavelet theory, functional analysis, and systems and control theory (see the monograph [2] for detail). This wide range of applications has motivated many authors to study the computation of joint spectral radius. Though even the approximation of joint spectral radius is in general an NP-hard problem [3], there are now a vast

*Corresponding author

Email addresses: masaki.ogura@ttu.edu (M. Ogura), clyde.f.martin@ttu.edu (C. F. Martin)

amount of efficient methods for the approximation of joint spectral radius [4, 5, 6] and also their implementations on mathematical softwares [7].

The result [4] by Blondel and Nesterov is of a particular theoretical interest because it characterizes joint spectral radius as the limit of another joint spectral characteristics called L^p -norm joint spectral radius when p tends to ∞ . Given a finite set $\mathcal{M} = \{A_1, \dots, A_N\}$ of real and square matrices of a fixed dimension and a parameter $p \geq 1$, the L^p -norm joint spectral radius (p -radius for short) of \mathcal{M} is defined by

$$\rho_{p,\mathcal{M}} := \lim_{k \rightarrow \infty} \left(N^{-k} \sum_{i_1, \dots, i_k \in \{1, \dots, N\}} \|A_{i_k} \cdots A_{i_1}\|^p \right)^{1/kp}, \quad (1)$$

where $\|\cdot\|$ denotes any matrix norm. Firstly introduced [8, 9] for $p = 1$ and then extended [10] for a general p , L^p -norm joint spectral radius has found many applications in various areas of applied mathematics (see [11] and references therein). In particular p -radius has an application to the stability theory of stochastic switched systems [12, 13, 14], which is a dynamical system whose structure randomly experiences abrupt changes [15, 16].

Recently this “original” version of L^p -norm joint spectral radius was extended to probability distributions [13]. Roughly speaking, the extension makes it possible to consider the p -radius of a set of *infinitely* many matrices and is useful when, for example, one wants to study the stability of a stochastic switched system with infinitely many subsystems that naturally arise as a result of uncertainty in modeling of dynamical systems. Being an extension, the p -radius of distributions inherits [13] from the p -radius of sets of matrices the characterization [10] as the spectral radius of a matrix. Though the characterization is valid only either when p is an even integer or when matrices in \mathcal{M} leave a common proper cone invariant, it still covers several interesting cases that appear in the stability analysis of stochastic switched linear systems. Then it is natural to expect that the other properties of the p -radius of sets of matrices can be extended to the p -radius of distributions.

In this paper we show that the characterization by Blondel and Nesterov [4] is still valid when we use the p -radius of probability distributions. This extension in particular circumvents the finiteness limitation of the original characterization. Since the proof for the original result relies on the finiteness of the number of matrices, it cannot be directly applied to the current setting. Instead, our proof extensively utilizes so-called cone linear absolute norms [17] and the approximation of a given set of possibly infinitely

many matrices by subsets having a certain uniformity property.

As a theoretical application of the characterization of joint spectral radius, we will discuss the stability of switched linear systems. We will present a novel characterization of the stability of a switched linear system for an arbitrary switching signal with a so-called stochastic Lyapunov function [18, 19, 20], which is a positive definite functional whose value decreases along the trajectory of the switched linear system in expectation. The characterization in particular deduces the existence of stochastic Lyapunov functions from stability and hence is a variant of the converse Lyapunov theorems [21, 22] in systems and control theory. The construction of stochastic Lyapunov functions is also investigated.

This paper is organized as follows. After preparing necessary notations in Section 2, in Section 3 we give a brief overview of the joint spectral radius of sets of matrices and the L^p -norm joint spectral radius of probability distributions. Then Section 4 gives the characterization of joint spectral radius as the limit of L^p -norm joint spectral radius. In Section 5 we discuss the application of the characterization to the stability theory of switched linear systems.

2. Mathematical Preliminaries

Let \mathbb{R}_+ denote the set of nonnegative real numbers. For $x \in \mathbb{R}^n$ its Euclidean norm is denoted by $\|x\|$, if not explicitly stated otherwise. For a real matrix A its maximal singular value is denoted by $\|A\|$. If A is square then its spectral radius is denoted by $\rho(A)$. When A is symmetric and negative semidefinite we write $A \preceq 0$. Let $\mathcal{M} \subset \mathbb{R}^{n \times n}$. The interior and the boundary of \mathcal{M} are denoted by $\text{int } \mathcal{M}$ and $\partial \mathcal{M}$, respectively. The distance between A and \mathcal{M} is defined by $d(A, \mathcal{M}) := \inf_{M \in \mathcal{M}} \|A - M\|$.

Let Ω be a probability space with a probability measure μ . The support of μ , denoted by $\text{supp } \mu$, is defined as the closed set such that $\mu((\text{supp } \mu)^c) = 0$ and, if G is open and $G \cap (\text{supp } \mu) \neq \emptyset$, then $\mu(G \cap \text{supp } \mu) > 0$. Dirac's delta distribution on $x \in \Omega$ is denoted by δ_x . For an integrable random variable X on Ω its expected value is denoted by $E[X]$.

2.1. Proper Cones

A subset $K \subset \mathbb{R}^n$ is called a cone if K is closed under multiplication by nonnegative numbers. The cone is said to be solid if it possesses a nonempty interior. We say that a cone is pointed if $x, -x \in K$ implies $x = 0$. We

say that K is proper if it is closed, convex, solid, and pointed. Let $K \subset \mathbb{R}^n$ be a proper cone. A matrix $A \in \mathbb{R}^{n \times n}$ is said to leave K invariant, written $A \geq^K 0$, if $AK \subset K$. The set of all real matrices leaving K invariant is denoted by $\pi(K)$ or simply by π . Let $B \in \mathbb{R}^{n \times n}$. By $A \geq^K B$ we mean $A - B \geq^K 0$. A set $\mathcal{M} \subset \mathbb{R}^{n \times n}$ is said to leave K invariant if any $A \in \mathcal{M}$ leaves K invariant. A is said to be K -positive if $A(K - \{0\}) \subset \text{int } K$ and we write $A >^K 0$. We understand $A >^K B$ in the obvious way. It is known that [24, p. 16]

$$\text{int } \pi = \{A \in \mathbb{R}^{n \times n} : A >^K 0\}. \quad (2)$$

Also we can show the next lemma.

Lemma 2.1. *The boundary $\partial\pi$ is a null set with respect to the Lebesgue measure.*

Proof. In general, the boundary of a convex set in $\mathbb{R}^{n \times n}$ is a null set with respect to the Lebesgue measure [23, Theorem 1]. This proves the claim because π is clearly convex. \square

A class of norms called cone linear absolute norms (see, e.g., [17]) plays an important role in this paper. A norm $\|\cdot\|$ on \mathbb{R}^n is said to be cone absolute [17] with respect to a proper cone K if, for every $x \in \mathbb{R}^n$,

$$\|x\| = \inf_{\substack{v, w \in K \\ x = v - w}} \|v + w\|. \quad (3)$$

Also we say that $\|\cdot\|$ is cone linear with respect to K if there exists f in the dual cone $K^* := \{f \in \mathbb{R}^n : f^\top x \geq 0 \text{ for every } x \in K\}$ such that

$$\|x\| = f^\top x \quad (4)$$

for every $x \in K$. A norm that is cone linear and cone absolute with respect to a proper cone is said to be cone linear absolute. It is known [17] that every $f \in \text{int}(K^*)$ yields a cone linear absolute norm $\|\cdot\|_f$ determined by (3) and (4). The norm $\|\cdot\|_f$ induces a norm on $\mathbb{R}^{n \times n}$ as

$$\|A\|_f := \sup_{x \in \mathbb{R}^n} \frac{\|Ax\|_f}{\|x\|_f}. \quad (5)$$

When f is irrelevant we simply denote $\|\cdot\|_f$ by $\|\cdot\|$. Some useful properties of this norm are quoted from [17] in the next lemma.

Lemma 2.2. *Let K be a proper cone and let $\|\cdot\|$ be a cone linear absolute norm with respect to K .*

1. *If $A \geq^K 0$ then*

$$\|A\| = \sup_{x \in K} \frac{\|Ax\|}{\|x\|}. \quad (6)$$

2. *If $A_i \geq^K B_i \geq^K 0$ for every $i = 1, \dots, k$ then*

$$\|A_k \cdots A_1\| \geq \|B_k \cdots B_1\|. \quad (7)$$

Proof. The first statement follows from [17, Theorem 2.1]. The second one is also proved in [17] when $k = 1$. Then the general case follows from the obvious relationship $A_k \cdots A_1 \geq^K B_k \cdots B_1 \geq^K 0$. \square

2.2. Lifts and Kronecker Products

Another notion that is used extensively in this paper is the lift of real vectors. Let $p \geq 1$ be an integer and let $x \in \mathbb{R}^n$. The p -lift (see, e.g., [5]) of x , denoted by $x^{[p]}$, is defined as the real vector of length $n_p = \binom{n+p-1}{p}$ with its elements being the lexicographically ordered monomials $\sqrt{\alpha!} x^\alpha$ indexed by all the possible exponents $\alpha = (\alpha_1, \dots, \alpha_n) \in \{0, 1, \dots, p\}^n$ such that $\alpha_1 + \dots + \alpha_n = p$, where $\alpha! := p! / (\alpha_1! \cdots \alpha_n!)$. For $A \in \mathbb{R}^{n \times n}$ we define the $n_p \times n_p$ matrix $A^{[p]}$ as the unique matrix [2] satisfying $(Ax)^{[p]} = A^{[p]}x^{[p]}$ for every $x \in \mathbb{R}^n$. For a subset \mathcal{M} of $\mathbb{R}^{n \times n}$ we define $\mathcal{M}^{[p]} = \{M^{[p]} : M \in \mathcal{M}\}$. Also for real matrices A and B , $A \otimes B$ denotes the Kronecker product [25] of A and B . Define the Kronecker power $A^{\otimes p}$ by $A^{\otimes 1} := A$ and $A^{\otimes(p)} = A^{\otimes(p-1)} \otimes A$ recursively for a general p . We define $\mathcal{M}^{\otimes p} := \{M^{\otimes p} : M \in \mathcal{M}\}$. It is known that if AB is defined then

$$(AB)^{\otimes p} = A^{\otimes p} B^{\otimes p}. \quad (8)$$

The next lemma collects some properties of p -lifts and Kronecker products proved in [4].

Lemma 2.3 ([4]). *Let $\mathcal{M} \subset \mathbb{R}^{n \times n}$.*

1. *$\mathcal{M}^{[2]}$ leaves a proper cone invariant.*
2. *If \mathcal{M} leaves a proper cone invariant then $\mathcal{M}^{\otimes p}$ also leaves a proper cone invariant for every $p \geq 1$.*

For a probability distribution μ on $\mathbb{R}^{n \times n}$ we define the probability distribution $\mu^{\otimes p}$ on $\mathbb{R}^{n^p \times n^p}$ as the image [26, Section 3.6] of μ under the measurable mapping $(\cdot)^{\otimes p}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n^p \times n^p}$. Let f be a measurable function on $\mathbb{R}^{n^p \times n^p}$. If A and B are independent random variables following μ and $\mu^{\otimes p}$, respectively, then we can show that

$$E[f(B)] = E[f(A^{\otimes p})]. \quad (9)$$

We also define $\mu^{[p]}$ as the image of μ under $(\cdot)^{[p]}: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n^p \times n^p}$.

3. Joint Spectral Characteristics

This section briefly overviews the notions of joint spectral radius and L^p -norm joint spectral radius. The joint spectral radius [2] of a bounded set $\mathcal{M} \subset \mathbb{R}^{n \times n}$ is defined by

$$\hat{\rho}(\mathcal{M}) := \limsup_{k \rightarrow \infty} \sup_{A_1, \dots, A_k \in \mathcal{M}} \|A_k \cdots A_1\|^{1/k}.$$

One of the important applications of joint spectral radius is in the stability theory of switched linear systems [15]. Define the switched linear system $\Sigma_{\mathcal{M}}$ by

$$\Sigma_{\mathcal{M}} : x(k+1) = A_k x(k), \quad A_k \in \mathcal{M} \quad (10)$$

where $x(0) = x_0 \in \mathbb{R}^n$ is a constant vector. We say that $\Sigma_{\mathcal{M}}$ is *absolutely exponentially stable* [27] if there exist $C > 0$ and $\gamma \in [0, 1)$ such that $\|x(k)\| \leq C\gamma^k \|x_0\|$ for every \mathcal{M} -valued sequence $\{A_k\}_{k=0}^{\infty}$ and x_0 . This stability is characterized by joint spectral radius as follows (see, e.g., [2]).

Proposition 3.1. $\Sigma_{\mathcal{M}}$ is absolutely exponentially stable if and only if $\hat{\rho}(\mathcal{M}) < 1$.

The following lemma lists some other properties of joint spectral radius. To state the lemma we recall that the set $\mathfrak{K}(\mathbb{R}^{n \times n})$ of compact and nonempty subsets of $\mathbb{R}^{n \times n}$ becomes a complete metric space [28] if it is endowed with the Hausdorff metric given by

$$H(\mathcal{M}, \mathcal{N}) := \max \left\{ \max_{A \in \mathcal{M}} d(A, \mathcal{N}), \max_{B \in \mathcal{N}} d(B, \mathcal{M}) \right\}. \quad (11)$$

Lemma 3.2. *The following statements are true.*

1. The restriction of the mapping $\hat{\rho}$ to the metric space $\mathfrak{R}(\mathbb{R}^{n \times n})$ is continuous [28, Lemma 3.5].
2. It holds [2, Proposition 2.5] that, for any $p \geq 1$,

$$\hat{\rho}(\mathcal{M}^{[p]}) = \hat{\rho}(\mathcal{M})^p. \quad (12)$$

Then we turn to the L^p -norm joint spectral radius of probability distributions introduced in [13]. Let μ be a probability distribution on $\mathbb{R}^{n \times n}$ and let A_k ($k = 1, 2, \dots$) be random variables independently following μ . Also let p be a positive integer. The L^p -norm joint spectral radius (p -radius for short) of μ is defined by

$$\rho_{p,\mu} := \lim_{k \rightarrow \infty} (E[\|A_k \cdots A_1\|^p])^{1/pk}. \quad (13)$$

This definition extends the L^p -norm joint spectral radius of a set of finitely many matrices shown in (1). One can check that if $\text{supp } \mu$ is bounded then $\rho_{p,\mu}$ exists and is finite [13]. Thus, without being explicitly stated, we assume that probability distributions appearing in this paper have a bounded support. Though in general the computation of p -radius is a difficult problem [11], the next simple formula for p -radius is available under certain assumptions.

Proposition 3.3 ([13, Theorem 2.5]). *Assume that one of the following conditions is true:*

- A₁. p is even;
- A₂. $\text{supp } \mu$ leaves a proper cone invariant.

Then

$$\rho_{p,\mu} = \rho(E[A^{\otimes p}])^{1/p},$$

where A is a random variable following μ .

We also quote from [13] other properties of p -radius that will be used in this paper.

Lemma 3.4. *Let $p \geq 1$ be arbitrary.*

1. If $p \leq q$ then $\rho_{p,\mu} \leq \rho_{q,\mu}$.

2. It holds that

$$\rho_{p,\mu} \geq \rho(E[A^{\otimes p}])^{1/p}. \quad (14)$$

3. For every $m \geq 1$,

$$\rho_{p,\mu^{\otimes m}} = \rho_{p,\mu^{[m]}} = \rho_{mp,\mu}^m. \quad (15)$$

Proof. The first two statements can be found in [13]. The last statement can be proved in the same way as (12). \square

Remark 3.5. By the equivalence of the norms on a finite dimensional vector space, the value of L^p -norm joint spectral radius is independent of the norm used in (13).

4. Limit Formula for Joint Spectral Radius

This section presents a novel limit formula for joint spectral radius. We state the next assumption on a probability distribution μ on $\mathbb{R}^{n \times n}$.

A₃. The singular part μ_s of μ is a linear combination of finitely many Dirac measures, i.e., either $\mu_s = 0$ or there exist positive numbers p_1, \dots, p_N and matrices M_1, \dots, M_N such that

$$\mu_s = p_1 \delta_{M_1} + \dots + p_N \delta_{M_N}. \quad (16)$$

Notice that any of the assumptions from A₁ to A₃ does not require $\text{supp } \mu$ to consist of only finitely many matrices. The next theorem is the main result of this paper.

Theorem 4.1. *Let μ be a probability distribution satisfying A₂ and A₃ and let $\mathcal{M} = \text{supp } \mu$. Then*

$$\hat{\rho}(\mathcal{M}) = \lim_{p \rightarrow \infty} \rho_{p,\mu}.$$

Proposition 3.3 allows us to state the theorem in the following equivalent form, which extends the limit formula of the joint spectral radius of a set of finitely many matrices given in [4].

Theorem 4.2. *Let μ be a probability distribution satisfying A₂ and A₃ and let $\mathcal{M} = \text{supp } \mu$. Then*

$$\hat{\rho}(\mathcal{M}) = \lim_{p \rightarrow \infty} \rho(E[A^{\otimes p}])^{1/p}.$$

If μ is the uniform distribution on a finite set then the theorem recovers [4, Theorem 3]. As a simple illustration of the present theorem let us see the next example.

Example 4.3. Let $\gamma > 0$ and let μ be the uniform distribution on $[0, \gamma]$. Clearly μ is absolutely continuous and $\mathcal{M} = \text{supp } \mu = [0, \gamma]$ leaves the proper cone \mathbb{R}_+ of \mathbb{R} invariant. It is easy to observe $\hat{\rho}(\mathcal{M}) = \gamma$ and $\rho(E[A^{\otimes p}]) = \gamma^p/(p+1)$. Therefore $\lim_{p \rightarrow \infty} \rho(E[A^{\otimes p}])^{1/p} = \gamma = \hat{\rho}(\mathcal{M})$, as expected. The characterization in [4] cannot be applied to this simple example as μ has an infinite support.

We can use the above limit formula to generalize another limit formula of joint spectral radius given in [29].

Corollary 4.4. *If μ is of the form (16) then*

$$\hat{\rho}(\mathcal{M}) = \lim_{p \rightarrow \infty} \rho(E[A^{\otimes(2p)}])^{1/(2p)} \quad (17)$$

$$= \limsup_{p \rightarrow \infty} \rho(E[A^{\otimes p}])^{1/p}. \quad (18)$$

It is remarked that, setting μ to be the uniform distribution in this corollary, we can recover Theorem 2.1 in [29].

Proof. We shall apply Theorem 4.1 to $\mu^{[2]}$, which satisfies both A_2 and A_3 because its support $\{A_1^{[2]}, \dots, A_N^{[2]}\}$ leaves a proper cone invariant by Lemma 2.3 and also $\mu^{[2]} = \sum_{i=1}^N p_i \delta_{A_i^{[2]}}$ is a finite sum of point masses. Therefore Theorem 4.1 shows $\hat{\rho}(\text{supp } \mu^{[2]}) = \lim_{p \rightarrow \infty} \rho_{p, \mu^{[2]}}$. This equation implies the first equality (17) because (12) shows $\hat{\rho}(\text{supp } \mu^{[2]}) = \hat{\rho}(\mathcal{M})^2$ and also (15) and Proposition 3.3 yield $\rho_{p, \mu^{[2]}} = \rho_{2p, \mu}^2 = \rho(E[A^{\otimes(2p)}])^{1/p}$. Then let us show the second equation (18). Using the inequality (14), the monotonicity of p -radius (Lemma 3.4), and Proposition 3.3, we can show

$$\begin{aligned} \rho(E[A^{\otimes(2p-1)}])^{1/(2p-1)} &\leq \rho_{2p-1, \mu} \\ &\leq \rho_{2p, \mu} \\ &= \rho(E[A^{\otimes(2p)}])^{1/(2p)}. \end{aligned}$$

This inequality and (17) prove the equation (18). \square

4.1. Proof

This section gives the proof of Theorem 4.1. Let μ be a probability distribution on $\mathbb{R}^{n \times n}$ with bounded support. First we observe that the definitions of p -radius and joint spectral radius immediately show $\hat{\rho}(\mathcal{M}) \geq \rho_{p,\mu}$. Since $\rho_{p,\mu}$ is non-decreasing with respect to p by Lemma 3.4, the limit $\lim_{p \rightarrow \infty} \rho_{p,\mu}$ exists and satisfies

$$\hat{\rho}(\mathcal{M}) \geq \lim_{p \rightarrow \infty} \rho_{p,\mu}. \quad (19)$$

Therefore, to prove Theorem 4.1, we need to show

$$\hat{\rho}(\mathcal{M}) \leq \lim_{p \rightarrow \infty} \rho_{p,\mu} \quad (20)$$

under the assumption that μ satisfies A_2 and A_3 . In the rest of this section we prove inequality (20). In the sequel it is assumed that μ satisfies both A_2 and A_3 . Let $\mathcal{M} = \text{supp } \mu$ and let K be the proper cone left invariant by \mathcal{M} .

We first note that showing (20) is not as straightforward as showing (19). Inequality (20) means that the maximum growth rate $\hat{\rho}(\mathcal{M})$ of the products of matrices from \mathcal{M} can be attained by the p th averaged growth rate $\rho_{p,\mu}$ of the products as $p \rightarrow \infty$. The difficulty in showing the inequality is that the set of sequences giving the maximum growth rate can have a very small probability, possibly zero, and therefore we should not expect that the rate can be captured by the averaged growth rate. For example, the joint spectral radius γ in Example 4.3 results from the singleton $\{\gamma\}$, which is a null set for μ .

We avoid the above mentioned problem by first focusing on well-behaving subsets of \mathcal{M} , and then approximating \mathcal{M} by a sequence of such subsets. For $M \in \mathbb{R}^{n \times n}$ let

$$\mathcal{S}_M := \{A \in \mathbb{R}^{n \times n} : A \geq^K M\}.$$

Then we define \mathfrak{M} as a family of measurable and nonempty subsets \mathcal{N} of \mathcal{M} satisfying the following property: for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$\mu(\mathcal{S}_{(1-\epsilon)M}) \geq \delta \quad (21)$$

for every $M \in \mathcal{N}$. The next proposition shows that the joint spectral radius of a subset belonging to \mathfrak{M} admits an estimate of the form (20). Roughly speaking, inequality (21) will be used to guarantee that μ is always “aware” of products of matrices with almost maximum growth rates. The uniformity of the lower bound δ with respect to M plays a key role.

Proposition 4.5. *If $\mathcal{N} \in \mathfrak{M}$ then $\hat{\rho}(\mathcal{N}) \leq \lim_{p \rightarrow \infty} \rho_{p,\mu}$.*

Proof. If $\hat{\rho}(\mathcal{N}) = 0$ then the inequality holds vacuously. Assume $\hat{\rho}(\mathcal{N}) > 0$. Then, without loss of generality we can assume $\hat{\rho}(\mathcal{N}) = 1$ by scaling matrices in \mathcal{N} by the factor $1/\hat{\rho}(\mathcal{N})$. Take a cone linear absolute norm $\|\cdot\|$ with respect to K . By Proposition 3.1, there exist $c > 0$ and $\{M_k\}_{k=1}^\infty \subset \mathcal{N}$ such that $\|M_k \cdots M_1\| > c$ for infinitely many k . Take an arbitrary $\gamma < 1$ and define $\epsilon := 1 - \gamma$. Let us take the corresponding $\delta > 0$ satisfying (21). Observe that if $A_i \in \mathcal{S}_{(1-\epsilon)M_i}$ then $A_i \geq^K (1 - \epsilon)M_i = \gamma M_i$ so that, by (7), we have $\|A_k \cdots A_1\| \geq \gamma^k \|M_k \cdots M_1\| > c\gamma^k$. Therefore

$$\mu^k(\{(A_1, \dots, A_k) : \|A_k \cdots A_1\| > c\gamma^k\}) \geq \prod_{i=1}^k \mu(\mathcal{S}_{(1-\epsilon)M_i}) \geq \delta^k$$

and hence, by Markov's inequality, we obtain $E[\|A_k \cdots A_1\|^p] > \delta^k (c\gamma^k)^p$, which implies $E[\|A_k \cdots A_1\|^p]^{1/kp} > \delta^{1/p} c^{1/k} \gamma$. Taking the limit $k \rightarrow \infty$ in this inequality shows $\rho_{p,\mu} \geq \delta^{1/p} \gamma$ by Remark 3.5. Thus we obtain $\lim_{p \rightarrow \infty} \rho_{p,\mu} \geq \gamma$. Since γ can be made arbitrarily close to 1 we see $\lim_{p \rightarrow \infty} \rho_{p,\mu} \geq 1 = \hat{\rho}(\mathcal{N})$, as desired. \square

If we could show that \mathcal{M} is in \mathfrak{M} then Proposition 4.5 proves inequality (20). In this paper, however, we leave open the problem of checking $\mathcal{M} \in \mathfrak{M}$ and take another approach via the approximation of \mathcal{M} by elements in \mathfrak{M} . Let us decompose μ as

$$\mu = \mu_c + \mu_s, \tag{22}$$

where μ_c is an absolutely continuous measure and μ_s is either the zero measure or is of the form (16). Clearly $\mathcal{M} = (\text{supp } \mu_c) \cup (\text{supp } \mu_s)$. For $r > 0$ we define

$$\pi_r := \{M \in \pi : d(M, \partial\pi) \geq r\}.$$

Notice that $\pi_r \subset \text{int } \pi$. Finally we let

$$\mathcal{M}_r := (\pi_r \cap \text{supp } \mu_c) \cup \text{supp } \mu_s.$$

We shall show that this \mathcal{M}_r is indeed in \mathfrak{M} and, furthermore, as $r \rightarrow 0$, $\hat{\rho}(\mathcal{M}_r)$ converges to $\hat{\rho}(\mathcal{M})$. Let us begin with the next observation.

Lemma 4.6. *There exists $r_0 > 0$ such that $\mathcal{M}_r \subset \mathfrak{K}(\mathbb{R}^{n \times n})$ for every $r < r_0$.*

Proof. Since \mathcal{M}_r is always compact, we need show that \mathcal{M}_r is not empty for every $r < r_0$ for some $r_0 > 0$. Since \mathcal{M}_r is decreasing with respect to r , it is sufficient to show that there exists $r_0 > 0$ such that \mathcal{M}_{r_0} is nonempty. Assume the contrary, i.e., $\mathcal{M}_r = \emptyset$ for every $r > 0$. Then it must be that $\text{supp } \mu_c \subset \partial\pi$ and $\mu_s = 0$. The latter condition shows that μ_c is nonzero. Thus the former condition shows that the nonzero and absolutely continuous measure μ_c is concentrated on a null set by Lemma 2.1, which is a contradiction. \square

In the sequel we assume $r < r_0$. In order to show $\mathcal{M}_r \in \mathfrak{M}$ we will need the next lemma.

Lemma 4.7. *Assume that μ is absolutely continuous (i.e., $\mu_s = 0$). Let $M \in \mathbb{R}^{n \times n}$ be arbitrary. If a sequence $\{M_k\}_{k=1}^\infty \subset \mathbb{R}^{n \times n}$ converges to M then*

$$\begin{aligned} \lim_{k \rightarrow \infty} \mu(\mathcal{S}_M \setminus \mathcal{S}_{M_k}) &= 0, \\ \lim_{k \rightarrow \infty} \mu(\mathcal{S}_{M_k} \setminus \mathcal{S}_M) &= 0. \end{aligned} \tag{23}$$

Proof. We only prove the first equation (23) because the second one can be proved in a similar way. By shifting the point M to the origin and also μ accordingly, without loss of generality we can assume that $M = 0$. In this case $\mathcal{S}_M = \pi$. Notice that the boundary $\partial\pi$ is a null set with respect to μ by Lemma 2.1. Therefore it is sufficient to show that, for every $A \in \mathcal{M} \setminus \partial\pi \subset \text{int } \pi$,

$$\lim_{k \rightarrow \infty} \chi_{\pi \setminus \mathcal{S}_{M_k}}(A) = 0, \tag{24}$$

where χ_S denotes the characteristic function for a set S . Since $A \in \text{int } \pi$ we have $A >^K 0$ by (2). Therefore the origin is in the open set $G = \{B \in \mathbb{R}^{n \times n} : B <^K A\} = A - \text{int } \pi$. Hence, since $\{M_k\}_{k=1}^\infty$ converges to 0, if k is sufficiently large then M_k is in G , i.e., $A >^K M_k$, which shows $A \in \mathcal{S}_{M_k}$. Thus (24) actually holds. \square

Then we can prove the next proposition.

Proposition 4.8. *If $0 < r < r_0$ then $\mathcal{M}_r \in \mathfrak{M}$.*

Proof. Fix $\epsilon > 0$ and $r > 0$. Define $\phi: \mathcal{M}_r \rightarrow \mathbb{R}$ by $\phi(M) := \mu(\mathcal{S}_{(1-\epsilon)M})$. First assume that μ is absolutely continuous with respect to the Lebesgue measure, i.e., $\mu_s = 0$. Since \mathcal{M}_r is compact, it is sufficient to show that ϕ is

continuous and positive. In order to show that ϕ is continuous at $M \in \mathcal{M}_r$, let $\{M_k\}_{k=1}^\infty$ be a sequence of \mathcal{M}_r converging to M . Then we can see that

$$|\phi(M) - \phi(M_k)| \leq \mu(\mathcal{S}_{(1-\epsilon)M} \setminus \mathcal{S}_{(1-\epsilon)M_k}) + \mu(\mathcal{S}_{(1-\epsilon)M_k} \setminus \mathcal{S}_{(1-\epsilon)M}),$$

which converges to 0 as $k \rightarrow \infty$ by Lemma 4.7. Therefore ϕ is continuous. Then let us show that $\phi(M) > 0$. Since $\mathcal{M}_r \subset \text{int } \pi$ we have $M >^K 0$ and therefore $M >^K (1-\epsilon)M$. This shows that the open set $\text{int } \mathcal{S}_{(1-\epsilon)M}$ and $\text{supp } \mu$ has a nonempty intersection containing M . Therefore $\mu(\text{int } \mathcal{S}_{(1-\epsilon)M}) > 0$ and hence $\phi(M) \geq \mu(\text{int } \mathcal{S}_{(1-\epsilon)M}) > 0$, as desired.

Then we consider the general case. Decompose μ as (22). On the one hand, from the above argument, there exists a constant $\delta_c > 0$ such that $\mu_c(\mathcal{S}_{(1-\epsilon)M}) \geq \delta_c$ for every $M \in \pi_r \cap \text{supp } \mu_c$. On the other hand, if $M \in \text{supp } \mu_s$ then $M = M_i \geq^K 0$ for some $1 \leq i \leq N$. Therefore $M_i \in \mathcal{S}_{(1-\epsilon)M}$ because $M_i \geq^K (1-\epsilon)M$. Hence $\phi(M) \geq \mu(\{M_i\}) = p_i > 0$. Thus we can see that $\delta = \min(\delta_c, p_1, \dots, p_N)$ is a desired constant. This argument is valid even when $\mu_c = 0$ and therefore completes the proof. \square

This proposition shows $\hat{\rho}(\mathcal{M}_r) \leq \lim_{p \rightarrow \infty} \rho_{p,\mu}$ for every $0 < r < r_0$ by Proposition 4.5. Therefore, if we could show

$$\hat{\rho}(\mathcal{M}) = \lim_{r \rightarrow 0} \hat{\rho}(\mathcal{M}_r) \quad (25)$$

then we can complete the proof of inequality (20). The rest of this section is devoted to the proof of (25). For the proof we will need the next lemma.

Lemma 4.9. *For every $A \in \mathcal{M}$ it holds that*

$$\lim_{r \rightarrow 0} d(A, \mathcal{M}_r) = 0. \quad (26)$$

Proof. Let $A \in \mathcal{M}$ be arbitrary. First we suppose $\mu_s = 0$. Let $\epsilon > 0$ be arbitrary and let \mathcal{B} denote the open ball in $\mathbb{R}^{n \times n}$ with center A and radius ϵ . Since $\mu(\mathcal{B}) > 0$ and the set $\partial\pi$ has the Lebesgue measure 0 by Lemma 2.1, $\mathcal{B} \cap \mathcal{M}$ is not contained in $\partial\pi$. Therefore we can take $M \in \mathcal{B} \cap \mathcal{M}$ that is not in $\partial\pi$. Let $\delta := d(M, \partial\pi) > 0$ and assume $r \leq \delta$. Then $M \in \mathcal{M}_\delta \subset \mathcal{M}_r$ and therefore $d(A, \mathcal{M}_r) \leq \|A - M\| < \epsilon$. This shows (26) since $\epsilon > 0$ was arbitrary.

Then let us consider the general case. If $A \in \text{supp } \mu_c$ then, since $\mathcal{M}_r \supset \pi_r \cap \text{supp } \mu_c$ we can see $\lim_{r \rightarrow 0} d(A, \mathcal{M}_r) \leq \lim_{r \rightarrow 0} d(A, \pi_r \cap \text{supp } \mu_c) = 0$,

where in the last equation we applied the above argument for the special case $\mu_s = 0$ to the normalized probability measure $\mu_c/(\mu_c(\mathbb{R}^{n \times n}))$. On the other hand, if $A \in \text{supp } \mu_s$ then (26) clearly holds because A is also in \mathcal{M}_r and hence $d(A, \mathcal{M}_r) = 0$ for every r . This completes the proof. \square

Now we can prove (25).

Proof of (25). By Lemma 3.2 it is sufficient to show $\lim_{r \rightarrow 0} H(\mathcal{M}, \mathcal{M}_r) = 0$. Notice that the distance $H(\mathcal{M}, \mathcal{M}_r)$ is well defined for each r . Since the set \mathcal{M}_r is decreasing with respect to r , the distance $H(\mathcal{M}, \mathcal{M}_r)$ is increasing with respect to r so that the limit $\lim_{r \rightarrow 0} H(\mathcal{M}, \mathcal{M}_r)$ does exist. We assume $\lim_{r \rightarrow 0} H(\mathcal{M}, \mathcal{M}_r) > 0$ to derive a contradiction. In this case there exists $\epsilon > 0$ such that $H(\mathcal{M}, \mathcal{M}_r) > \epsilon$ for every $r > 0$. By the definition of the Hausdorff metric (11) this implies $\max_{A \in \mathcal{M}} d(A, \mathcal{M}_r) > \epsilon$ because $\mathcal{M}_r \subset \mathcal{M}$. Therefore there exists $A_r \in \mathcal{M}$ such that $d(A_r, \mathcal{M}_r) > \epsilon$ for each $r > 0$. Now let $\{r_i\}_{i=1}^\infty$ be a positive sequence decreasingly converging to 0. Since \mathcal{M} is compact, there exists a subsequence $\{r'_i\}_{i=1}^\infty$ of $\{r_i\}_{i=1}^\infty$ such that $\{A_{r'_i}\}_{i=1}^\infty$ converges to some $A \in \mathcal{M}$. Using the triangle inequality we can show

$$\begin{aligned} d(A, \mathcal{M}_{r'_i}) &\geq d(A_{r'_i}, \mathcal{M}_{r'_i}) - d(A_{r'_i}, A) \\ &> \epsilon - d(A_{r'_i}, A) \end{aligned}$$

and hence $\limsup_{r \rightarrow 0} d(A, \mathcal{M}_r) \geq \epsilon$, which contradicts to (26). \square

5. Lyapunov Theorem for Switched Linear Systems

As a theoretical application of the limit formulas obtained in the last section, in this section we show a novel characterization of the absolute exponential stability of the switched linear system $\Sigma_{\mathcal{M}}$ defined by (10) via so-called stochastic Lyapunov functions. We will also investigate the construction of stochastic Lyapunov functions. Define the stochastic switched linear system Σ_{μ} by

$$\Sigma_{\mu} : x(k+1) = A_k x(k), \quad A_k \text{ follows } \mu \text{ independently}$$

where $x(0) = x_0 \in \mathbb{R}^n$ is a constant. We say that Σ_{μ} is *exponentially stable in p th mean* (*p th mean stable* for short) if there exist $C > 0$ and $\gamma \in [0, 1)$ such that

$$E[\|x(k)\|^p] \leq C \gamma^{pk} \|x_0\|^p$$

for every $x_0 \in \mathbb{R}^n$. We call γ as a *growth rate of the p th mean*. As is expected, p th mean stability is closely related to p -radius.

Proposition 5.1 ([13]). Σ_μ is p th mean stable if and only if $\rho_{p,\mu} < 1$.

Now we introduce the notion of stochastic Lyapunov functions for Σ_μ . The following definition extends the ones in [18, 19, 20] by allowing us to study p th mean stability for a general p .

Definition 5.2. We say that $V: \mathbb{R}^n \rightarrow \mathbb{R}$ is a *stochastic Lyapunov function of degree p* for Σ_μ if there exist positive numbers C_1, C_2 such that

$$C_1\|x\|^p \leq V(x) \leq C_2\|x\|^p \quad (27)$$

and $\gamma \in [0, 1)$ such that

$$E[V(Ax)] \leq \gamma^p V(x) \quad (28)$$

for every $x \in \mathbb{R}^n$, where A is a random variable following μ . We say that V has a *growth rate* γ .

The next theorem is the main result of this section and provides a connection between the stability of deterministic switched linear systems and that of stochastic switched linear systems.

Theorem 5.3. Let μ be a probability distribution satisfying A_2 and A_3 . Let $\mathcal{M} = \text{supp } \mu$. Then $\Sigma_{\mathcal{M}}$ is absolutely exponentially stable if and only if there exists $\gamma < 1$ such that, for every $p \geq 1$, Σ_μ admits a stochastic Lyapunov function of degree p with growth rate at most γ .

Remark 5.4. In contrast to the well known characterization of absolute exponential stability with the existence of a single Lyapunov function called a common Lyapunov function [22], Theorem 5.3 characterizes absolute exponential stability with the existence of *infinitely many* stochastic Lyapunov functions. Also we notice that the above theorem deduces the existence of Lyapunov functions from the absolute exponential stability and hence can be considered as a version of converse Lyapunov theorems [22, 21] in the systems and control theory literature.

For the proof of Theorem 5.3 we will need the next proposition.

Proposition 5.5. Let μ be a probability distribution on $\mathbb{R}^{n \times n}$.

1. If Σ_μ admits a stochastic Lyapunov function with degree p and growth rate $\gamma < 1$ then Σ_μ is p th mean stable with growth rate γ .

2. If Σ_μ is p th mean stable then, for every $\gamma \in (\rho_{p,\mu}, 1)$, Σ_μ admits a stochastic Lyapunov function with degree p and growth rate γ .

Proof. First assume that Σ_μ admits a stochastic Lyapunov function V with degree p and growth rate $\gamma < 1$. Let $x_0 \in \mathbb{R}^n$ be arbitrary. Using an induction we can show $E[V(x(k))] \leq \gamma^{pk} V(x_0)$. Therefore (27) shows $E[\|x(k)\|^p] \leq (C_2/C_1) \gamma^{pk} \|x_0\|^p$. Thus Σ_μ is p th mean stable with growth rate γ .

On the other hand assume that Σ_μ is p th mean stable and let $\gamma \in (\rho_{p,\mu}, 1)$ be arbitrary. We follow the construction of Lyapunov functions in [30]. Define $h_k := E[\|A_k \cdots A_1\|^p]^{1/pk}$, where A_1, A_2, \dots are random variables following μ independently. Since $h_k \rightarrow \rho_{p,\mu}$ as $k \rightarrow \infty$, there exists k_0 such that $h_{k_0} \leq \gamma$. Define $V: \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$V(x) := \sum_{k=0}^{k_0-1} \frac{E[\|A_k \cdots A_1 x\|^p]}{\gamma^{pk}}, \quad (29)$$

where, if $k = 0$, the product $A_k \cdots A_1$ is understood to be the identity matrix with probability one. Let us show that V is a stochastic Lyapunov function for Σ_μ with degree p and growth rate γ . It is immediate to see that $\|x\|^p \leq V(x) \leq [\sum_{k=0}^{k_0-1} (h_k/\gamma)^{pk}] \|x\|^p$, where the first inequality can be obtained by truncating the series in (29) at $k = 0$. Moreover the independence of the random variables A_k yields

$$\begin{aligned} E[V(Ax)] &= \sum_{k=0}^{k_0-1} \frac{E[\|A_{k+1} A_k \cdots A_1 x\|^p]}{\gamma^{pk}} \\ &= \gamma^p \sum_{k=1}^{k_0} \frac{E[\|A_k \cdots A_1 x\|^p]}{\gamma^{pk}}. \end{aligned} \quad (30)$$

Since the last term of this sum can be bounded as

$$\frac{E[\|A_{k_0} \cdots A_1 x\|^p]}{\gamma^{pk_0}} \leq \frac{h_{k_0}^{pk_0} \|x\|^p}{\gamma^{pk_0}} \leq \|x\|^p,$$

the equation (30) shows $E[V(Ax)] \leq \gamma^p V(x)$. This completes the proof of the proposition. \square

Remark 5.6. When μ is the uniform distribution on a finite set, Proposition 5.5 reduces to [31, Proposition 1], where for its proof the author makes use of so-called extremal norms.

Now we prove Theorem 5.3.

Proof of Theorem 5.3. Assume that $\Sigma_{\mathcal{M}}$ is absolutely exponentially stable. Then there exist $C > 0$ and $0 \leq \gamma' < 1$ such that $\|x(k)\| < C\gamma'^k\|x_0\|$ for any choice of the sequence $\{A_k\}_{k=0}^\infty \subset \mathcal{M}$ and x_0 . Now let $\gamma \in (\gamma', 1)$ be arbitrary and let us fix $p \geq 1$. Since $E[\|x(k)\|^p] < C^p\gamma'^{pk}$, the system Σ_μ is p th mean stable with growth rate γ' . Therefore, by Proposition 5.5, Σ_μ admits a stochastic Lyapunov function with degree p and growth rate γ .

On the other hand assume that there exists $\gamma < 1$ such that, for every $p \geq 1$, Σ_μ admits a stochastic Lyapunov function of degree p with growth rate γ . By Proposition 5.5, Σ_μ is p th mean stable with growth rate γ , which furthermore implies $\rho_{p,\mu} \leq \gamma$ by Proposition 5.1. Therefore $\lim_{p \rightarrow \infty} \rho_{p,\mu} \leq \gamma < 1$. Hence, by Theorem 4.1, we obtain $\hat{\rho}(\mathcal{M}) < 1$, which gives the absolute exponential stability of $\Sigma_{\mathcal{M}}$ by Proposition 3.1. \square

5.1. Construction of Stochastic Lyapunov Functions

The realization (29) of a stochastic Lyapunov function as a sum involving several expected values of products of matrices is not useful in practice. In this section we will show that, if either the conditions A_1 or A_2 in Proposition 3.3 holds, then we can construct stochastic Lyapunov functions easily.

The next theorem covers the case when A_1 holds.

Theorem 5.7. *Assume that Σ_μ is p th mean stable and A_1 holds. Let $q := p/2$ and let $\gamma \in (\rho_{p,\mu}, 1)$ be arbitrary. Then the function*

$$V(x) = (x^{\otimes q})^\top H x^{\otimes q}, \quad (31)$$

where the positive definite matrix $H \in \mathbb{R}^{n^q \times n^q}$ is a solution of the linear matrix inequality

$$E[(A^{\otimes q})^\top H A^{\otimes q}] - \gamma^p H \preceq 0, \quad (32)$$

is a stochastic Lyapunov function for Σ_μ of degree p with growth rate γ .

To prove this theorem we need the following special case of the theorem given in [13].

Proposition 5.8. *Assume that Σ_μ is mean square stable. Let $\gamma \in (\rho_{2,\mu}, 1)$ be arbitrary. Then the function $V(x) = x^\top H x$ on \mathbb{R}^n , where the positive definite matrix $H \in \mathbb{R}^{n \times n}$ is a solution of the linear matrix inequality $E[A^\top H A] - \gamma^2 H \preceq 0$, is a stochastic Lyapunov function for Σ_μ with degree 2 and growth rate γ .*

Let us prove Theorem 5.7.

Proof of Theorem 5.7. Assume that Σ_μ is p th mean stable and let $\gamma \in (\rho_{p,\mu}, 1)$ be arbitrary. Since (15) and Proposition 5.1 show $\rho_{2,\mu^{\otimes q}} = \rho_{p,\mu}^q < 1$, the system $\Sigma_{\mu^{\otimes q}}$ is mean square stable again by Proposition 5.1. Since $\gamma^q > \rho_{p,\mu}^q = \rho_{2,\mu^{\otimes q}}$, Proposition 5.8 implies that Σ_μ admits a stochastic Lyapunov function $W(x) = (Hx, x)$ on \mathbb{R}^{n^q} with growth rate γ^q . Moreover the matrix H can be obtained as a solution of the matrix linear inequality $E[B^\top HB] - (\gamma^q)^2 H \preceq 0$, where B is a random variable following $\mu^{\otimes q}$. This linear matrix inequality is indeed equivalent to (32). Now define $V: \mathbb{R}^n \rightarrow \mathbb{R}$ by (31) or, equivalently, by $V(x) := W(x^{\otimes q})$. Let us show that V is a stochastic Lyapunov function of degree p with growth rate γ . Using (8) and (9) we can see that

$$\begin{aligned} E[V(Ax)] &= E[W(A^{\otimes q} x^{\otimes q})] \\ &= E[W(Bx^{\otimes q})] \\ &\leq (\gamma^q)^2 W(x^{\otimes q}) \\ &= \gamma^p V(x). \end{aligned}$$

To show that an inequality of the form (27) holds for V , notice that there exist positive constants C_1, C_2 satisfying $C_1 \|y\|^2 \leq W(y) \leq C_2 \|y\|^2$ for every $y \in \mathbb{R}^{n^q}$ because H is positive definite. Letting $y = x^{\otimes q}$ we obtain (27) by the well-known identity $\|x^{\otimes q}\| = \|x\|^q$ (see, e.g., [25]) that holds for a general q and $x \in \mathbb{R}^n$ provided $\|\cdot\|$ is the Euclidean norm. Hence V is a stochastic Lyapunov function of degree p with growth rate γ . \square

Then we consider the condition A_2 . In order to proceed we here quote a basic result on K -positive matrices from [32].

Lemma 5.9 ([32, Theorem 4.4]). *Let K be a proper cone and assume $A >^K 0$.*

1. *A has a simple eigenvalue $\rho(A)$, which is greater than the magnitude of any other eigenvalue of A ;*
2. *The eigenvector corresponding to the eigenvalue $\rho(A)$ is in $\text{int } K$.*

Then we prove the next proposition. Recall that, for a proper cone K and $f \in \text{int}(K^*)$ the matrix norm $\|\cdot\|_f$ is defined by (5), which is induced by the cone linear absolute norm $\|\cdot\|_f$ satisfying (3) and (4).

Proposition 5.10. *Let $K \subset \mathbb{R}^n$ be a proper cone and assume that $M \geq^K 0$. Also let $\epsilon > 0$ be arbitrary. Then there exists $f \in \text{int}(K^*)$ such that $\|M\|_f < \rho(M) + \epsilon$.*

Proof. First assume $M >^K 0$. By Lemma 5.9 the matrix M admits the Jordan canonical form $J = V^{-1}MV$ where $V \in \mathbb{R}^{n \times n}$ is an invertible matrix whose columns are the generalized eigenvectors of M and $J \in \mathbb{R}^{n \times n}$ is of the form

$$J = \begin{bmatrix} J_0 & 0 \\ 0 & \rho(M) \end{bmatrix}$$

for some upper diagonal matrix $J_0 \in \mathbb{R}^{(n-1) \times (n-1)}$. Define $f \in \mathbb{R}^n$ by

$$V^{-1} = \begin{bmatrix} * \\ f^\top \end{bmatrix}.$$

Then we can easily see that f is an eigenvector of M^\top corresponding to the eigenvalue $\rho(M)$. Since K^* is proper and M^\top is K^* -positive (see, e.g., [24]), Lemma 5.9 shows $f \in \text{int}(K^*)$. Also since $f^\top Mx = \rho(M)f^\top x$, the equation (6) shows $\|M\|_f = \rho(M)$.

Then we consider the general case of $M \geq^K 0$. Let $\epsilon > 0$ be arbitrary and take an arbitrary $P >^K 0$. Then there exists $\delta > 0$ such that $\rho(M + \delta P) < \rho(M) + \epsilon$ because $\rho(M + \delta P) \rightarrow \rho(M)$ as $\delta \rightarrow 0$ by the continuity of spectral radius. Since $M + \delta P >^K 0$, the above argument shows that there exists $f \in K^*$ satisfying $\|M + \delta P\|_f = \rho(M + \delta P) < \rho(M) + \epsilon$. Finally, since $0 \leq^K M \leq^K M + \delta P$, Lemma 2.2 shows $\|M\|_f \leq \|M + \delta P\|_f$ and thus we obtain the desired inequality. \square

The next theorem enables us to construct a stochastic Lyapunov function when A_2 holds.

Theorem 5.11. *Assume that Σ_μ is p th mean stable and A_2 holds. Let $\gamma \in (\rho_{p,\mu}, 1)$ be arbitrary. Then there exists a cone linear absolute norm $\|\cdot\|$ on \mathbb{R}^{n^p} such that $V(x) = \|x^{\otimes p}\|$ is a stochastic Lyapunov function for Σ_μ with degree p and growth rate γ .*

Proof. Assume that Σ_μ is p th mean stable and let $\gamma \in (\rho_{p,\mu}, 1)$ be arbitrary. Let K be a proper cone left invariant by $\text{supp } \mu$.

First we consider the special case $p = 1$. Since A leaves K invariant with probability one we have $E[A] \geq^K 0$. Also Proposition 3.3 shows

$\rho(E[A]) = \rho_{1,\mu} < \gamma$. Therefore, by Proposition 5.10, there exists a cone linear absolute norm $\|\cdot\|_f$ on \mathbb{R}^n such that $\|E[A]\|_f < \gamma$. Let us show that $V(x) = \|x\|_f$ gives a stochastic Lyapunov function for Σ_μ with degree 1 and growth rate γ .

The inequality of the form (27) clearly holds for some positive constants C_1 and C_2 by the equivalence of the norms on a finite dimensional normed vector space. To show (28) let $x \in \mathbb{R}^n$ and $\delta > 0$ be arbitrary. Since $\|\cdot\|_f$ is cone linear absolute there exist $x_1, x_2 \in K$ such that $x = x_1 - x_2$ and $\|x_1\|_f + \|x_2\|_f = \|x_1 + x_2\|_f \leq \|x\|_f + \delta$. Moreover we have $Ax_i \in K$ and therefore $\|Ax_i\|_f = f^\top Ax_i$ with probability one. Thus it follows that

$$\begin{aligned} E[\|Ax_i\|_f] &= f^\top E[A]x_i \\ &= \|E[A]x_i\|_f \\ &< \gamma\|x_i\|_f \end{aligned}$$

for each $i = 1, 2$. Hence, since $\|Ax\|_f = \|Ax_1 - Ax_2\|_f \leq \|Ax_1\|_f + \|Ax_2\|_f$,

$$\begin{aligned} E[\|Ax\|_f] &< \gamma(\|x_1\|_f + \|x_2\|_f) \\ &\leq \gamma(\|x\|_f + \delta). \end{aligned}$$

Since $\delta > 0$ was arbitrary we obtain $E[\|Ax\|_f] \leq \gamma\|x\|_f$. This inequality shows that, since $x \in \mathbb{R}^n$ was arbitrary, $\|\cdot\|_f$ is a stochastic Lyapunov function for Σ_μ with growth rate γ and degree 1.

Then let us give the proof for a general p . Since $\rho_{1,\mu^{\otimes p}} = \rho_{p,\mu}^p < 1$ by (15), $\Sigma_{\mu^{\otimes p}}$ is first mean stable by Proposition 5.1. Also notice that, by Lemma 2.3, $\text{supp}(\mu^{\otimes p}) = (\text{supp } \mu)^{\otimes p}$ leaves a proper cone in \mathbb{R}^{n^p} , say K_p , invariant. Thus, by the above result for $p = 1$, since $\gamma^p > \rho_{p,\mu}^p = \rho_{1,\mu^{\otimes p}}$, the system $\Sigma_{\mu^{\otimes p}}$ admits a stochastic Lyapunov function $\|\cdot\|_g$ with growth rate γ^p and degree 1, where $\|\cdot\|_g$ is a cone linear absolute norm on \mathbb{R}^{n^p} with respect K_p . Now we define $V: \mathbb{R}^n \rightarrow \mathbb{R}$ by $V(x) := \|x^{\otimes p}\|_g$. Then, in the same way as the proof of Theorem 5.7, we can show that V is a stochastic Lyapunov function for Σ_μ with degree p and growth rate γ . \square

Example 5.12. Consider the probability distribution

$$\mu = \begin{bmatrix} [0, 1.5] & [0, 1.8] \\ [0, 0.15] & [0, 1.2] \end{bmatrix},$$

where each interval denotes the uniform distribution on it. Clearly $\text{supp } \mu$ leaves the proper cone \mathbb{R}_+^2 invariant and moreover we can see $\rho(E[A]) < 1$.

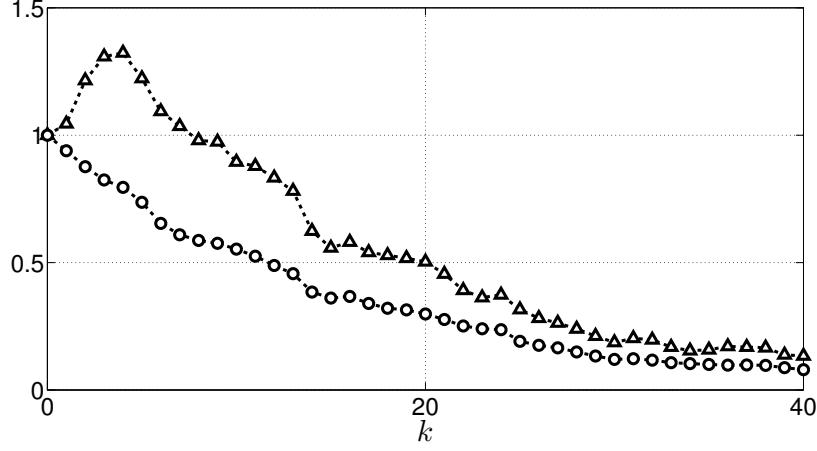


Figure 1: The sample means of the Lyapunov function (circle) and the Euclidean norm (triangle)

Therefore Propositions 5.1 and 3.3 show that Σ_μ is first mean stable and hence, by Proposition 5.5, Σ_μ admits a stochastic Lyapunov function of degree 1. Following the proof of Theorem 5.11 we find a stochastic Lyapunov function $\|x\|_f$ for Σ_μ where $f = [0.3838 \quad 1]^\top$. We generate 200 sample paths of Σ_μ with the initial state $x_0 = [0 \quad 1]^\top$. Figure 1 shows the sample means of the stochastic Lyapunov function $\|x(k)\|_f$ and the Euclidean norm $\|x(k)\|$. We can see that the sample mean of the Lyapunov function decreases at the most of time instances, while that of the Euclidean norm shows an oscillating behavior. It is remarked that the sample mean of the Lyapunov function is not necessarily decreasing because it is actually different from the expectation. Taking more sample paths in general makes the sample mean closer to the expectation by the law of large numbers and therefore is more likely to yield a decreasing sample mean. Figure 2 shows the average of the sample paths and the contour plot of the stochastic Lyapunov function and the Euclidean norm. The figure graphically illustrates that the sample mean is evolving in such a way that the value of constructed Lyapunov function almost decreases.

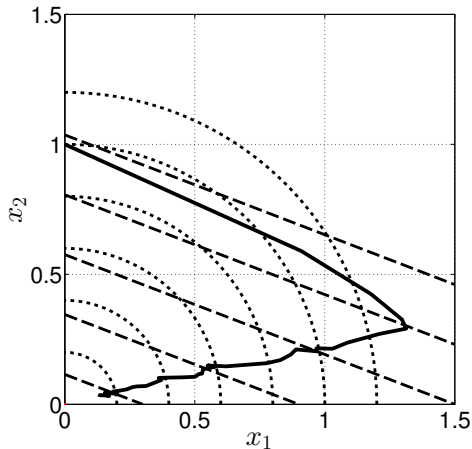


Figure 2: The averaged sample path (solid) and the level plots of the Lyapunov function (dashed) and the Euclidean norm (dotted)

6. Conclusion and Discussion

This paper presented a characterization of the joint spectral radius of a set of matrices as the limit of the L^p -norm joint spectral radius of a probability distribution on the set when $p \rightarrow \infty$ under the assumption that the distribution has a certain regularity and a support leaving a proper cone invariant. The obtained characterization extends the ones in the literature by allowing the set to have infinitely many matrices. Based on the result, we also presented a novel characterization of the absolute exponential stability of switched linear systems via the existence of stochastic Lyapunov functions of any higher degrees. The construction of stochastic Lyapunov functions is also studied.

Understanding the behavior of p -radius as $p \rightarrow 0$ is one of the problems closely related to the problem studied in this paper. It is known [33] that, as $p \rightarrow 0$, the p -radius converges to so-called Lyapunov exponent [34] of random products of matrices, which is known to characterize so-called almost sure stability of stochastic switched systems. However the characterization in [33] is proved under the assumption that the number of matrices in the set from which one takes a matrix is finite. It would be interesting to investigate if one can allow the number of matrices to be infinite with the approach taken in this paper.

References

- [1] G.-C. Rota, W. G. Strang, A note on the joint spectral radius, *Indag. Math* 22 (1960) 379–381.
- [2] R. M. Jungers, The Joint Spectral Radius, volume 385 of *Lecture Notes in Control and Information Sciences*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [3] J. N. Tsitsiklis, V. D. Blondel, The Lyapunov exponent and joint spectral radius of pairs of matrices are hard – when not impossible – to compute and to approximate, *Mathematics of Control, Signals, and Systems* 10 (1997) 31–40.
- [4] V. D. Blondel, Y. Nesterov, Computationally efficient approximations of the joint spectral radius, *SIAM Journal on Matrix Analysis and Applications* 27 (2005) 256–272.
- [5] P. A. Parrilo, A. Jadbabaie, Approximation of the joint spectral radius using sum of squares, *Linear Algebra and its Applications* 428 (2008) 2385–2402.
- [6] V. Y. Protasov, R. M. Jungers, V. D. Blondel, Joint spectral characteristics of matrices: a conic programming approach, *SIAM Journal on Matrix Analysis and Applications* 31 (2010) 2146–2162.
- [7] G. Vankeerberghen, J. Hendrickx, R. M. Jungers, JSR: A toolbox to compute the joint spectral radius, in: *Proceedings of the 17th international conference on Hybrid systems: computation and control*, pp. 151–156.
- [8] R.-Q. Jia, Subdivision schemes in L_p spaces, *Advances in Computational Mathematics* 3 (1995) 309–341.
- [9] Y. Wang, Two-scale dilation equations and the mean spectral radius, *Random and Computational Dynamics* 4 (1996) 49–72.
- [10] V. Y. Protasov, The generalized joint spectral radius. A geometric approach, *Izvestiya: Mathematics* 61 (1997) 995–1030.

- [11] R. M. Jungers, V. Y. Protasov, Fast methods for computing the p -radius of matrices, *SIAM Journal on Scientific Computing* 33 (2011) 1246–1266.
- [12] R. M. Jungers, V. Y. Protasov, Weak stability of switching dynamical systems and fast computation of the p -radius of matrices, in: 49th IEEE Conference on Decision and Control, pp. 7328–7333.
- [13] M. Ogura, C. F. Martin, Generalized joint spectral radius and stability of switching systems, *Linear Algebra and its Applications* 439 (2013) 2222–2239.
- [14] M. Ogura, C. F. Martin, On the mean stability of a class of switched linear systems, in: 52nd IEEE Conference on Decision and Control, pp. 97–102.
- [15] H. Lin, P. Antsaklis, Stability and stabilizability of switched linear systems: a survey of recent results, *IEEE Transactions on Automatic Control* 54 (2009) 308–322.
- [16] R. Shorten, F. Wirth, O. Mason, K. Wulff, C. King, Stability criteria for switched and hybrid systems, *SIAM Review* 49 (2007) 545–592.
- [17] T. I. Seidman, H. Schneider, M. Arav, Comparison theorems using general cones for norms of iteration matrices, *Linear Algebra and its Applications* 399 (2005) 169–186.
- [18] J. Bertram, P. Sarachik, Stability of circuits with randomly time-varying parameters, *IRE Transactions on Circuit Theory* 6 (1959) 260–270.
- [19] G. Ahmadi, On the mean square stability of linear difference equations, *Applied Mathematics and Computation* 241 (1979) 233–241.
- [20] X. Feng, K. A. Loparo, Y. Ji, H. J. Chizeck, Stochastic stability properties of jump linear systems, *IEEE Transactions on Automatic Control* 31 (1992) 38–53.
- [21] W. Dayawansa, C. Martin, A converse Lyapunov theorem for a class of dynamical systems which undergo switching, *IEEE Transactions on Automatic Control* 44 (1999) 751–760.

- [22] A. Molchanov, Y. Pyatnitskiy, Criteria of asymptotic stability of differential and difference inclusions encountered in control theory, *Systems & Control Letters* 13 (1989) 59–64.
- [23] R. Lang, A note on the measurability of convex sets, *Archiv der Mathematik* 47 (1986) 90–92.
- [24] A. Berman, R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, 1979.
- [25] J. Brewer, Kronecker products and matrix calculus in system theory, *IEEE Transactions on Circuits and Systems* 25 (1978) 772–781.
- [26] V. I. Bogachev, *Measure Theory*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.
- [27] L. Gurvits, Stability of discrete linear inclusion, *Linear Algebra and its Applications* 231 (1995) 47–85.
- [28] F. Wirth, The generalized spectral radius and extremal norms, *Linear Algebra and its Applications* 342 (2002) 17–40.
- [29] J. Xu, M. Xiao, A characterization of the generalized spectral radius with Kronecker powers, *Automatica* 47 (2011) 1530–1533.
- [30] T. Ando, M.-H. Shih, Simultaneous contractibility, *SIAM Journal on Matrix Analysis and Applications* 19 (1998) 487–498.
- [31] V. Y. Protasov, Extremal L_p -norms of linear operators and self-similar functions, *Linear Algebra and its Applications* 428 (2008) 2339–2356.
- [32] J. S. Vandergraft, Spectral properties of matrices which have invariant cones, *SIAM Journal on Applied Mathematics* 16 (1968) 1208–1222.
- [33] Y. Fang, K. A. Loparo, On the relationship between the sample path and moment Lyapunov exponents for jump linear systems, *IEEE Transactions on Automatic Control* 47 (2002) 1556–1560.
- [34] L. Arnold, A formula connecting sample and moment stability of linear stochastic systems, *SIAM Journal on Applied Mathematics* 44 (1984) 793–802.